Preparation and characterization of the RNase H domain of Moloney murine leukemia virus reverse transcriptase.

AUTHOR(S):

Nishimura, Kosaku; Yokokawa, Kanta; Hisayoshi, Tetsuro; Fukatsu, Kosuke; Kuze, Ikumi; Konishi, Atsushi; Mikami, Bunzo; Kojima, Kenji; Yasukawa, Kiyoshi

CITATION:

Nishimura, Kosaku ...[et al]. Preparation and characterization of the RNase H domain of Moloney murine leukemia virus reverse transcriptase.. Protein expression and purification 2015, 113: 44-50

ISSUE DATE:

2015-09

URL:

http://hdl.handle.net/2433/200910

Protein Expression and Purification

Preparation and characterization of the RNase H domain of Moloney murine leukemia virus reverse transcriptase

Kosaku Nishimura[a], Kanta Yokokawa[a], Tetsuro Hisayoshi[a], Kosuke Fukatsu[a], Ikumi Kuze[a], Atsushi Konishi[a], Bunzo Mikami[b], Kenji Kojima[a], Kiyoshi Yasukawa[a,*]

[a]*Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan*

[b]*Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan*

***** Corresponding author. Fax: +81-75-753-6265.

E-mail address: yasukawa@kais.kyoto-u.ac.jp (K. Yasukawa)

Abbreviations used: HIV-1, human immunodeficiency virus type 1; MMLV, Moloney murine leukemia virus; RNase H, ribonuclease H; RT, reverse transcriptase; PAGE, polyacrylamide gel electrophoresis

Moloney murine leukemia virus reverse transcriptase (MMLV RT) contains fingers, palm, thumb, and connection subdomains as well as an RNase H domain. The DNA polymerase active site resides in the palm subdomain, and the RNase H active site is located in the RNase H domain. The RNase H domain contains a positively charged α-helix called the C helix (H^{594}GEIYRRR^{601}), that is thought to be involved in substrate recognition. In this study, we expressed three versions of the RNase H domain in *Escherichia coli*, the wild-type domain (WT) (residues Ile498−Leu671) and two variants that lack the regions containing the C helix (Ile593−Leu603 and Gly595−Thr605, which we called ΔC1 and ΔC2, respectively) with a strep-tag at the N-terminus and a deca-histidine tag at the C-terminus. These peptides were purified from the cells by anion-exchange, Ni^{2+} affinity, and Strep-Tactin affinity column chromatography, and then the tags were removed by proteolysis. In an RNase H assay using a 25-bp RNA-DNA heteroduplex, WT, ΔC1, and ΔC2 produced RNA fragments ranging from 7 to 16 nucleotides (nt) whereas the full-length MMLV RT (Thr24−Leu671) produced 14−20-nt RNA fragments, suggesting that elimination of the fingers, palm, thumb, and connection subdomains affects the binding of the RNase H domain to the RNA-DNA heteroduplex. The activity levels of WT, ΔC1, and ΔC2 were estimated to be 1%, 0.01%, and 0.01% of full-length MMLV RT activity, indicating that the C helix is important, but not critical, for the activity of the isolated RNase H domain.

Keywords: Moloney murine leukemia virus; reverse transcriptase; RNase H activity; template-primer; thermostabilization

Introduction

Retroviral reverse transcriptase (RT) possesses RNA- and DNA-dependent DNA polymerase as well as RNase H activities. Moloney murine leukemia virus (MMLV) RT is a 75-kDa monomer, comprised of the fingers (Thr24−Asp124 and Phe156−Ser195), palm (Ile125−Phe155 and Pro196−Glu275), thumb (Gly276−Leu338), and connection (Pro339−Asp468) subdomains and an RNase H domain (Arg469−Leu671). The active site of the DNA polymerase activity is located in the palm subdomain while the RNase H activity is in the RNase H domain [1, 2].

Due to its high catalytic activity and fidelity, MMLV RT is extensively used in cDNA synthesis [3]. Thermostability of DNA polymerase activity of MMLV RT increases by inactivating the RNase H activity by the mutation of the catalytically important residue, Asp524 [4-6]. The same strategy was successful to improve the thermostabilities of RTs from human immunodeficiency virus type-1 (HIV-1) [7] and avian myeloblastosis virus (AMV) [5, 8].

The whole structure of MMLV RT has not been determined yet. From the crystal of full-length MMLV RT (Thr24−Leu671), the structures of the fingers, palm, thumb, and connection subdomains (Thr24−Asn479) have been determined, but that of the RNase H domain has not [1, 9]. On the other hand, from the crystal of the isolated RNase H domain, its whole structure has been determined [10]. In that study, the RNase H domain variant lacking the polypeptide Ile593-Leu603 was successfully crystallized, while the wild-type RNase H domain was not [10]. The deleted polypeptide region, Ile593-Leu603, contains a positively charged α-helix called the C helix (H^{594}GEIYRRR^{601}). In xenotropic murine leukemia virus-related virus (XMRV) RT, which has 95% amino-acid sequence homology

with MMLV RT, like MMLV RT, the structure of the isolated RNase H domain has been determined whereas the whole structure has not [11-13]. Under these backgrounds, Das *et al*. [1] and Cote *et al*. [2] have proposed the structural model of MMLV RT that the RNase H domain is positioned far from the fingers/palm/thumb domain. However, Pandey *et al*. have proposed another model that the polypeptide region Pro480-Arg506 of the RNase H domain, the structure of which has not been determined from the crystal of the full-length molecule or the RNase H domain, interacts with the fingers/palm/thumb subdomain, supporting the floor of the DNA polymerase active-site cleft. In MMLV RT, it is difficult to clarify the interaction between the DNA polymerase and RNase H active sites due to the lack of the crystal structure of the whole MMLV RT molecule.

The expression and purification of the isolated RNase H domains of MMLV RT in *Esherichia coli* were previolusly reported [1, 14-16]. They were expressed without tags [14, 16] or *S*-transferase fusion protein [1, 15]. In this study, we describe expression, purification, and characterization of three versions of the RNase H domain of MMLV RT, the wild-type domain and the two C helix-deficient variants. They were expressed with N-terminal strep-tag and C-terminal deca-histidine tag in *E. coli*, purified from the cells, and then the tags were removed by proteolysis. The enzyme preparations thus obtained were characterized and compared with full-length MMLV RT.

Materials and methods

Materials

[γ-^{32}P]dATP (111 TBq/mmol) was purchased from PerkinElmer (Waltham, MI).

Oligonucleotides were from Fasmak (Atsugi, Japan). The RT concentration was determined by the method of Bradford [17] using Protein Assay CBB Solution (Nacalai Tesque, Kyoto, Japan) with bovine serum albumin (Nacalai Tesque) as standard.

Plasmids

We previously constructed the expression plasmid for C-teminally $(His)_6$-tagged MMLV RT (pET-MRT) by inserting the C-teminally $(His)_6$-tagged MMLV RT gene into the *Nde*I and *Eco*RI sites of pET-22b(+) plasmid (Merck Bioscience, Tokyo, Japan). The expression plasmids for the wild-type RNase H domain (Ile498−Leu671) (WT), its variant lacking Ile593-Leu603 (ΔC1), and its variant lacking Gly595−Thr605 (ΔC2), each with C-teminal $(His)_6$ tag (pET22b-RNaseH-WT, pET22b-RNaseH-ΔC1, and pET22b-RNaseH-ΔC2, respectively) were constructed as follows. Site-directed mutagenesis was carried out using the following nucleotide primers 5'-CAACACAACTGCCATATGATCCTGGCCGAA-3' (497Nde) and 5'-TTCGGCCAGGATCATATGGCAGTTGTGTTG-3' (497Nde_cp) to change the nucleotide sequence CTTGAT encoding ^{496}Leu^{497}Asp into a *Nde*I recognition sequence CATATG. The resuting plasmid, pET-MRT(Nde), was digested with *Nde*I and self-ligated to produce the expression plasmid for WT, pET22b-RNaseH-WT. The 391-bp fragment (F1) was amplified from pET22b-RNaseH-WT using the following nucleotide primers 5'-TAATACGACTCACTATAGGG-3' (Pro) and 5'-GCCTTCTGATGTGAGATGGGCAGTAGCAAA-3' (DelI593-L603rev-4), and the 247-bp fragment (F2) was amplified using the following nucleotide primers 5'-TTTGCTACTGCCCATCTCACATCAGAAGGC-3' (DelI593-L603for-4) and 5'-

CTGAATTCTAGTGGTGATGGTGGTGGTGGAGGAGGGTAGAGGTGT-3' (MRT-HISB). The 608-bp fragment was amplified using the nucleotide primers Pro and MRT-HISB from the mixture of F1 and F2, digested with the restriction enzymes *Xba*I and *Eco*RI, and inserted in pET-22b(+) digested with *Xba*I and *Eco*RI to produce the expression plasmid for ΔC1, pET22b-RNaseH-ΔC1. The expression plasmid for ΔC2, pET22b-RNaseH-ΔC1 was constructed by the same method but using the nucleotide primers 5'-CTCTTTGCCTTCTGATGTGAGATGGGCAGT-3' (DelG595-T605rev-4) and 5'-ACTGCCCATATCCATTCAGAAGGCAAAGAG-3' (DelG595-T605for-4) instead of DelI593-L603rev-4 and DelI593-L603for-4, respectively.

The expression plasmids for WT, ΔC1, and ΔC2, each with an N-terminal strep tag and a C-teminal $(His)_{10}$ tag (pET52b-RNaseH-WT, pET52b-RNaseH-ΔC1, and pET52b-RNaseH-ΔC2, respectively) were constructed as follows. The 550-bp fragment was amplified with the following nucleotide primers 5'-TTTTGGTACCTTGATGTCCTGGCCGAAGCC-3' (497Kpn) and 5'-TTTTGAGCTCGAGGAGGGTAGAGGTGTC-3' (671Sac) from pET22b-RNaseH-WT, and the 517-bp fragments were amplified from 497Kpn and 671Sac from pET22b-RNaseH-ΔC1 or pET22b-RNaseH-ΔC2. These three fragments were digested with the restriction enzymes *Kpn*I and *Sac*I, and inserted in pET-52b(+) (Merck Bioscience) to produce pET52b-RNaseH-WT, pET52b-RNaseH-ΔC1, and pET52b-RNaseH-ΔC2, respectively.

Expression and purification of recombinant RNase H domain

E. coli strain BL21(DE3) [*F*-, *ompT*, $hsdS_B(r_B\text{-}m_B\text{-})$ *gal dcm* (DE3)] was transformed

with either pET52b-RNaseH-WT, pET52b-RNaseH-ΔC1, or pET52b-RNaseH-ΔC2. The overnight culture of the transformants (20 ml) was added to 2,000 ml of LB broth and in a 2-liter flask and incubated at 37°C under vigorous aeration by air-pump. When OD_{660} reached 0.6–0.8, 0.5 mL of 0.5 M IPTG was added and growth was continued at 30ºC for 4 h. After centrifugation at 10,000 × *g* for 10 min, the cells were harvested, suspended with 20 ml of 0.02 M potassium phosphate (pH 7.2), 2.0 mM dithiothreitol (DTT), 10% glycerol (buffer A) containing 1 mM phenylmethylsulfonyl fluoride (PMSF) and disrupted by sonication. After centrifugation at 20,000 × *g* for 40 min, the supernatant was collected and applied to a column [25 mm (inner diameter) × 120 mm] packed with Toyopearl DEAE-650M gel (Tosoh, Tokyo, Japan), previoulsy equilibrated with buffer A. The bound RNase H domain was eluted with buffer A containing 100 mM NaCl and applied to a HisTrap column, previously equilibrated with 50 mM Tris-HCl (pH 8.3), 200 mM KCl, 2.0 mM DTT, 10% glycerol (buffer B). The bound RNase H domain was eluted with buffer B containing 250 mM imidazole and applied to a Strep-Tactin Superflow column (IBA, Göttingen, Germany), previously equilibrated with 100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA (buffer C). The bound RNase H domain was eluted with buffer C containing 2.5 mM desthiobiotin, 50% glycerol. Two mg of the obtained RNase H domain was subjected to the digestion by 10 units of HRV 3C protease (Accelagen, San Diego, CA) to cleave the strept tag and 50 units of thrombin (GE Healthcare) to cleave the $(His)_{10}$ tag in 12.5 ml of 40 mM Tris-HCl (pH 7.5), 120 mM NaCl, 10% glycerol at 4°C for 36 h (buffer D). Then, the solution was three-fold diluted by adding 25 ml of water and further incubated at 4°C for 36 h. The digest was applied to a HisTrap column, previously equilibrated with buffer B. The RNase H domain without tags had weak binding ability to a HisTrap column, and it was collected from a HisTrap

column by eluting with buffer B containing 50 mM imidazole. In this buffer condition, the RNase H domain with deca-histidine tag was not eluted. The obtained RNase H domain without tags was applied to a PD-10 column packed with a Sephadex G-25 (GE Healthcare), previously equilibrated with 20 mM Tris-HCl (pH 7.0), containing 200 mM NaCl, 2.0 mM DTT, and 10% glycerol. The column was washed and eluted with the same buffer. Purified RNase H domain was stored at -80ºC before use.

SDS-PAGE

SDS-PAGE of the RNase H domain was carried out in a 12.5% polyacrylamide gel under reducing conditions. Proteins (1.2 μg) were reduced by treatment with 2.5% (v/v) of 2-mercaptoethanol at 100ºC for 10 min, and then applied onto the gel. A constant current of 40 mA was applied for 40 min. After electrophoresis, proteins were stained with Coomassie Brilliant Blue R-250. The molecular mass marker kit consisting of rabbit muscle phosphorylase B (97.2 kDa), bovine serum albumin (66.4 kDa), hen egg white ovalbumin (44.3 kDa), and bovine carbonic anhydrase (29.0 kDa) was purchased from Takara Bio Inc (Otsu, Japan).

CD spectroscopy

The CD spectra of the RNase H domain was measured with a 2-mm cell using a J-820 spectropolarimeter (Jasco, Tokyo, Japan) under the following condition: spectral range 200-250 nm; 25ºC; 100 nm min^{-1} scan rate; and 10 accumulations.

Fluorescence-based RNase H assay

Fluorescence-based RNase H activity assay was carried out according to the method by Parniak *et al* [18]. Briefly, an RNA/DNA hybrid (named R18/D18) was prepared by mixing the 3'-fluorescein modified 18-nucleotide (nt) RNA 5'-GAUCUGAGCCUGGGAGCU-3' (R18) and 5'-dabcyl-modified complementary 18-nt DNA (D18) with a molar ratio of 1.0:1.2 in 50 mM Tris-HCl buffer (pH 8.0) containing 60 mM KCl followed by the incubation at room temperature for 30 min. The RNase H reaction was carried out in 50 mM Tris-HCl buffer (pH 8.0) containing 60 mM KCl, 5.0 mM $MgCl_2$, 0.25 μM R18/D18 (the concentration is expressed as that of R18), and indicated concentrations of the RNase H domain or the full-length MMLV RT at 37ºC. An aliquot (800 μl) was taken from the reaction mixture at a predetermined time and added to 0.5 M EDTA (160 μl). The fluorescence spectra of the reaction products were measured with a Shimadzu RF-5300PC fluorescence spectrophotometer (Shimadzu, Kyoto, Japan) under the following conditions: excitation wavelength, 490 nm; spectral range 500-600 nm; 25ºC; 120 nm min^{-1} scan rate; and 3 accumulations. In each measurement, the control baseline was obtained with the corresponding buffer in the absence of protein.

Radioisotope-based RNase H assay

Radioisotope-based RNase H activity assay was carried out according to the method by Álvarez *et al* [19]. Briefly, an RNA/DNA hybrid (named R25/D25) was prepared by mixing 5'-[^{32}P]-labelled 25-nt RNA 5'-AUGUAUAGCCCUACCAGCAUUCTGG-3'

(R25) and unlabelled complementary 25-nt DNA (D25) by the method above described. The RNase H reaction (40 µl) was carried out in 50 mM Tris–HCl (pH 8.0) buffer, containing 50 mM KCl, 0.1 mM $MgCl_2$, 20 nM R25/D25, and indicated concentrations of the RNase H domain or the full-length MMLV RT at 37ºC. An aliquot (4 µl) was taken from the reaction mixture at a predetermined time and added to 4 µl of sample-loading buffer (10 mM EDTA, 90% (v/v) formamide, 3 mg/ml xylene cyanol FF, 3 mg/ml bromophenol blue, and 50 µM 31D/21D. (31D/21D was prepared by mixing unlabelled 31-nt DNA 5'-TTTTTTTTTAGGATACATATGGTTAAAGTAT-3' (D31) and unlabelled complementary complementary 21-nt DNA 5'-ATACTTTAACCATATGTATCC-3'(D21) by the method above described and used as a carrier DNA.) The RNA fragments of the reaction products were analyzed by denaturing 20% polyacrylamide gel electrophoresis followed by image scanning with a BAS-2500 scanner (Fujifilm, Tokyo, Japan) using the program Multi Gauge version 2.2 (Fujifilm).

Results

Expression and purification of the RNase H domain

We initially attempted to express the wild-type RNase H domain (WT) (Ile498−Leu671) and the two C helix-deficient variants lacking Ile593-Leu603 (ΔC1) and Gly595-Thr605 (ΔC2), with C-terminal $(His)_6$ tags at the C-terminus. We constructed three expression plasmids, pET22b-RNaseH-WT, pET22b-RNaseH-ΔC1, and pET22b-RNaseH-ΔC2 that contain the T7 promoter, an ATG initiation codon, a 540- or 507-bp sequence encoding 180 or 169 amino acids comprising the intact or C helix-deficient

RNase H domain (Ile498−Leu671), and a $(His)_6$ tag, and a stop codon. SDS-PAGE analysis of the soluble fractions of the *E. coli* transformants showed a 20-kDa protein band corresponding to WT or a 19-kDa protein band corresponding to ΔC1 and ΔC2 (data not shown). However, the RNase H domain could not be purified to homogeneity because several contaminating proteins, presumably C-terminal $(His)_6$-tagged RNase H domain degradation products, could not be completely removed (data not shown).

Next we attempted to express WT, ΔC1, and ΔC2 with both an N-terminal strep tag and a C-terminal $(His)_{10}$ tag. We constructed three expression plasmids, pET52b-RNaseH-WT, pET52b-RNaseH-ΔC1, and pET52b-RNaseH-ΔC2 (Fig. 1), that contain the T7 promoter, an ATG initiation codon, a 711- or 678-bp sequences encoding 237 or 206 amino acids comprising a strep-tag, an HRV 3C protease recognition site, the intact or C helix-deficient RNase H domain, a thrombin recognition site, and a $(His)_{10}$ tag, and a stop codon. SDS-PAGE analysis of the soluble fractions showed a 25-kDa protein band corresponding to WT in the transformants with pET52b-RNaseH-WT (lane 2 in Fig. 2A) and a 24-kDa protein band corresponding to ΔC1 or ΔC2 in the transformants with pET52b-RNaseH-ΔC1 or pET52b-RNaseH-ΔC2, respectively (data not shown).

To purify WT, ΔC1, and ΔC2, we used anion-exchange chromatography as the first step, the Ni^{2+}-sepharose chromatography as the second step, and the Strep-Tactin affinity chromatography as the third step. At each step, each fraction was tested for the presence of the 25-kDa or 24-kDa protein band by SDS-PAGE, and the active fractions were pooled based on coomassie staining. The SDS-PAGE analysis showed that the active fractions from the WT purification exhibited 25-kDa protein bands (Fig. 2A), and the purified preparations of WT, ΔC1, and ΔC2 exhibited 25-kDa or 24-kDa protein bands (Fig. 2B). In Fig. 2B, the 50 or 55 kDa faint bands were also observed; however their

orgins were unknown. Then, the N-terminal strep tag and C-terminal $(His)_{10}$ tag were removed by successive proteolysis with HRV 3C protease and thrombin, respectively. After proteolysis, the WT preparation exhibited a single 22-kDa protein band (Fig. 2C), the ΔC1 and ΔC2 preparations exhibited single 21-kDa protein bands (data not shown), and the protein bands corresponding to the 50 or 55 kDa faint bands obserbed in Fig. 2B were not observed (Fig. 2C for WT and and data not shown for ΔC1 and ΔC2). From the 2 L cultures, 0.5−1.0 mg of purified WT, ΔC1, and ΔC2 were obtained. In 20 mM Tris-HCl (pH 7.0), 200 mM NaCl, 2.0 mM DTT, and 10% glycerol, WT, ΔC1, and ΔC2 with tags could be concentrated to 1−2 mg/ml, whereas without tags, these proteins could be concentrated to 10 mg/ml (data not shown), indicating that the removing of the tags increased the solubility.

To examine if the deletion of Ile593-Leu603 or Gly595-Thr605 affects structural change of the isolated RNase H domain, we made spectroscopic analyses. UV spectroscopy showed that purified WT, ΔC1, and ΔC2 exhibited spectra with a deep trough at approximately 250 nm and a peak at 275 nm (Fig. 3B for WT and data not shown for ΔC1 and ΔC2). CD spectroscopy showed that purified WT, ΔC1, and ΔC2 exhibited negative ellipticities at 202–250 nm with peaks at approximately 208 nm and 222 nm (Fig. 3B for WT and data not shown for ΔC1 and ΔC2). No appreciable changes were observed in UV and CD) spectra, suggesting that ΔC1, and ΔC2 did not suffer from any global or drastic structural changes by the deletion.

Characterization of the RNase H domain

To analyze the RNase H activity of the isolated RNase H domains, we performed

fluorescence-based RNase H assay. An RNA/DNA hybrid (called R18/D18) consisting of a 3'-fluorescein modified 18-nt RNA (R18) and a 5'-dabcyl-modified 18-nt DNA (D18) was used as the substrate (Fig. 4A). R18/D18 is designed to emit strong fluorescence when R18 is cleaved at a site close to the 3' end, and the fluorescein-labeled RNA fragment dissociates from the complementary DNA strand. Figure 4B shows the difference in the fluorescence spectra of R18/D18 before and after the reaction. Based on the previous report, the $MgCl_2$ concentration was set at 5.0 mM [19]. When R18/D18 was incubated without the full-length RT or the isolated RNase H domain for 30 min, there was no change in the fluorescence spectra. When R18/D18 was incubated with 50 nM HIV-1 RT or 1.8 μM WT for 30 min, the fluorescence increased with the peak at 515 nm. When R18/D18 was incubated with ΔC1 or ΔC2 for 30 min, no fluorescence increase was observed. These results suggest that WT has RNase H activity and ΔC1 and ΔC2 do not.

To further analyze the RNase H activity of the isolated RNase H domains, we used a radioisotope-based RNase H assay. An RNA/DNA hybrid (R25/D25) consisting of a 5'-[^{32}P]-labelled 25-nt RNA (R25) and an unlabeled complementary DNA (D25) was used as the substrate (Fig. 5A). After the RNase H reaction, the RNA reaction products were analyzed by denaturing PAGE. First we examined the effects of $MgCl_2$ concentration on the RNase H activity of WT and found that reactions containing 0.1−0.2 mM yielded relatively higher amounts of products than those containing other concentrations (0−15 mM) (data not shown). Thus, 0.1 mM was used in subsequent experiments.

Figure 5B shows the denatured PAGE analysis of the products obtained from the reaction under various conditions. Unreacted R25/D25 showed only a single 25-nt RNA band. When R25/D25 was incubated with 30 nM full-length MMLV RT, 16−20-nt bands were detected at 15 s, and 14−17-nt bands were detected at 10 min. When R25/D25 was

incubated with 30 nM WT, only a 25-nt band was detected at both 15 s and 10 min. When R25/D25 was incubated with 1.0 or 10 μM WT, only a 25-nt band was detected at 15 s, whereas various bands (mainly 7−16-nt bands) were detected at 10 min. This indicated that the cleavage pattern of WT was different from that of the full-length MMLV RT and that the activity of WT was considerably weaker than that of the full-length MMLV RT. When R25/D25 was incubated with ΔC1 or ΔC2, various bands (mainly 7−16-nt bands) were only detected at a 10 μM and 10 min. Under all other reaction conditions tested, only the uncleaved 25-nt band was detected. This indicated that although the activities of ΔC1 and ΔC2 were weaker than that of WT, the cleavage patterns were the same as that of WT.

Figure 5C compares the time courses of the cleavage patterns obtained in the reactions with 30 nM full-length MMLV RT, 1 μM WT, and 10 μM ΔC1 and ΔC2. The results of a densitometry trace of each band were shown in Fig. S1. In full-length MMLV RT, 10−20-nt bands were detected, in which the amount of the 17-nt band was the highest at 5-30 min and that of the 14-nt band was the highest at 60 and 120 min. In WT, ΔC1, and ΔC2, 7−20-nt bands were detected, in which the amount of the 7- or 10-nt band was higher than those of other bands. The rates of the decrease in the 25-nt bands in the reaction with WT, ΔC1, and ΔC2 were 30%, 3%, and 3% of full-length MMLV RT, respectively. Based on these enzyme concentrations (30 nM for full-length MMLV RT, 1 μM for WT, and 10 μM for ΔC1, and ΔC2), the activities of WT, ΔC1, and ΔC2 were roughly estimated to be 1%, 0.01%, and 0.01% of full-length MMLV RT activity, respectively.

Discussion

We expressed the isolated RNase H domain of MMLV RT with N-terminal strep-tag and C-terminal deca-histidine tag. The use of two tags has an advantage as it allows the removal of degraded products using one of the tags. In our initial attempt to express the RNase H domain with only a C-terminal $(His)_6$-tag, the degraded products could not be completely removed, as described in the Results section. In addition, the removing of the tags was important because it increased the solubility. Based on the SDS-PAGE (Fig. 2) and spectroscopic analysis (Fig. 3) results, we considered the purity of the RNase H preparations to be sufficient for use in structural and functional analysis.

The DNA in the 25-bp RNA/DNA heteroduplex used in the radioisotope-based RNase H assay lacks the primer 3'-terminus (Fig. 5A). The reaction with full-length MMLV RT produced 17−20-nt RNA fragments (Fig. 5B), indicating that the position at which the RNA strand of the RNA/DNA hybrid is cleaved is 17−20 bp upstream of the 3'-terminus of the primer. This is the same position as previoulsly reported in assays of full-length MMLV RT using RNA/DNA hybrids with primer 3'-termini [20, 21]. This suggests that the DNA polymerase active site binds to the RNA/DNA hybrid in a similar manner whether it has the primer 3'-terminus or not.

Unlike the full-length MMLV RT, WT, ΔC1, and ΔC2 produced RNA fragments mainly ranging from 7 to 16 nt in length (Fig. 5B). We speculate a possible mechanism for the generation of fragments of these lengths as follows: For the full-length MMLV RT, the position at which the RNase H active site binds to the RNA strand of an RNA/DNA hybrid is predetermined, whereas for the isolated RNase H domains, the active site binding position is not; thus the cleavage site by the isolated RNase H domain

is different from that by the full-length MMLV RT. This suggests that elimination of the fingers, palm, thumb, and connection subdomains affects the binding of the RNase H domain to an RNA/DNA heteroduplex.

In this study, we expressed WT as well as ΔC1 and ΔC2 because the crystal structure of MMLV RT ΔC1 has been determined [1] and the structures of XMRV RT WT and ΔC2 have been determined [22, 23]. In the RNase H assay using a fluorescent substrate, WT exhibited activity, whereas ΔC1 and ΔC2 did not (Fig. 4). However, in the RNase H assay using a radioactive substrate, all three exhibited the activity (Fig. 5). Therefore, we concluded that the C helix is not critical for the activity of the isolated MMLV RT RNase H domain. However, the RNase H activities of ΔC1 and ΔC2 were only 10% of WT activity (Fig. 5C), suggesting that the C helix is important for activity. Similar results have been reported for full-length MMLV RT, and the C helix has been shown to be important for both DNA polymerase and RNase H activities [22, 23].

Unlike the MMLV RT RNase H domain, the RNase H domain of HIV-1 RT does not contain the C helix, and the isolated HIV-1 RT RNase H domain lacks activity [24-27]. Interestingly, insertion of a basic loop near at the active site of the isolated RNase H domain generated weak activity in the presence of Mn^{2+} [28]. Recently, Permanasari *et al.* reported that addition of bacterial RNase HI or HII substrate binding domain to the isolated RNase H domain also resulted in weak activity in the presence of Mn^{2+} [29]. These results supports the notion that the C helix is important for RNase H activity.

E. coli RNase HI contains the C helix [30] whereas the RNase H domains of MMLV RT and HIV-1 RT are C helix-deficient. Comparison of the structure of ΔC1 (PDB accession codes SHB5) with that of *E. coli* RNase HI (3AA2) suggests that the C helix of the RNase H domain of MMLV RT is positioned similary to that of *E. coli* RNase HI.

Comparison of the structure of ΔC1 with a complex containing the isolated HIV-1 RT RNase H domain and an RNA/DNA heteroduplex (1HYS) suggests that the C helix in the RNase H domain of MMLV RT interacts with the RNA strand of the RNA/DNA heteroduplex. Therefore, the C helix is thought to contribute to substrate recognition, as reported previously by Nowotny *et al.* [31].

In conclusion, we expressed the RNase H domain of MMLV RT in *E. coli* and purified it to homogeneity. Characterization of the recombinant proteins suggested that the C helix is involved in the substrate recognition and that the fingers, palm, thumb, and connection subdomains are involved in positioning of the cleavage sites in the RNA strand of an RNA/DNA heteroduplex. These recombinant RNase H domain proteins might be suitable for use in structural and functional analyses to elucidate the mechanisms of thermostabilization of MMLV RT through site-directed mutagenesis. An extensive study is currently underway to prepare crystals of WT, ΔC1, and ΔC2, with and without a mutation of the catalytic Asp (Asp524).

Acknowledgments

This study was supported in part (K.Y.) by Grants-in-Aid for Scientific Research (no. 21580110) from the Japan Society for the Promotion of Science and the Salt Science Researh Foundation.

References

[1] D. Das, M.M. Georgiadis, The crystal structure of the monomeric reverse

transcriptase from Moloney murine leukemia virus, Structure 12 (2004) 819-829.

[2] M.L. Coté, M.J. Roth, Murine leukemia virus reverse transcriptase: structural comparison with HIV-1 reverse transcriptase, Virus Res. 134 (2008) 186-202.

[3] A.R. Kimmel, S.L. Berger, Preparation of cDNA and the generation of cDNA libraries: overview, Methods Enzymol. 152 (1987) 307-316.

[4] M.L. Kotewicz, J.M. D'Alessio, K.M. Driftmier, K.P. Blodgett, G.F. Gerard, Cloning and overexpression of Moloney murine leukemia virus reverse transcriptase in *Escherichia coli*, Gene 35 (1985) 249-258.

[5] G.F. Gerard, R.J. Potter, M.D. Smith, K. Rosenthal, G. Dhariwal, J. Lee, D.K. Chatterjee, The role of template-primer in protection of reverse transcriptase from thermal inactivation, Nucleic Acids Res. 30 (2002) 3118-3129.

[6] M. Mizuno, K. Yasukawa, K. Inouye, Insight into the mechanism of the stabilization of Moloney murine leukaemia virus reverse transcriptase by eliminating RNase H activity, Biosci. Biotechnol. Biochem. 74 (2010) 440-442.

[7] K. Nishimura, M. Shinomura, A. Konishi, K. Yasukawa, Stabilization of human immunodeficiency virus type 1 reverse transcriptase by site-directed mutagenesis, Biotechnol. Lett. 35 (2013) 2165-2175.

[8] A. Konishi, K. Yasukawa, K. Inouye, Improving the thermal stability of avian myeloblastosis virus reverse transcriptase α-subunit by site-directed mutagenesis, Biotechnol. Lett. 34 (2012) 1209-1215.

[9] D. Lim, G.G. Gregorio, C. Bingman, E. Martinez-Hackert, W.A. Hendrickson, S.P. Goff, Crystal structure of the Moloney murine leukemia virus RNase H domain, J. Virol. 80 (2006) 8379-8389.

[10] E. Nowak, W. Potrzebowski, P.V. Konarev, J.W. Rausch, M.K. Bona, D.I. Svergun,

J.M. Bujnicki, S.F.J. Le Grice, M. Nowotny, Structural analysis of monomeric retroviral reverse transcriptase in complex with an RNA/DNA hybrid, Nucleic Acid Res. 41 (2013) 3874-3887.

[11] D. Zhou, S. Chung, M. Miller, S.F.J. Le Grice, A. Wlodawer, Crystal structures of the reverse transcriptase-associated ribonuclease H domain of xenotropic murine leukemia-virus related virus, J. Struct. Biol. 177 (2012) 638-645.

[12] J.H. Kim, S. Kang, S.-K. Jung, K.R. Yu, S.J. Chung, B.H. Chung, R.L. Erikson, B.Y. Kim, S.J. Kim, Crystal structure of xenotropic murine leukaemia virus-related virus (XMRV) ribonulease H, Biosci. Rep. 32 (2012) 455-463.

[13] V. Barrioluengo, Y. Wang, S.F.J. Le Grice, L. Menéndez-Arias, Intrinsic DNA synthesis fidelity of xenotropic murine leukemia virus-related virus reverse transcriptase, FEBS J. 279 (2012) 1433-1444.

[14] N. Tanese, S.P. Goff, Domain structure of the Moloney murine leukemia virus reverse transcriptases: mutational analysis and separate expression of the DNA poluymerase and RNase H activities, Proc. Natl. Acad. Sci. USA 85 (1988) 1777-1781.

[15] S.J. Schultz and J.J. Champoux, RNase H domain of Moloney murine leukemia virus reverse transcriptase retains activity byt requires the polymerase domain for specificity, J. Virol. 70 (1996) 8630-8638.

[16] E.R. Gooden and S. Marqusee, Folding the ribonulcease H domain of Moleney murine leukemia virus reverse transcriptase requires mtal binding or a short N-terminal extension, Proteins 33 (1998) 135-143.

[17] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem.

72 (1976) 248-254.

[18] M.A. Parniak, K.L. Min, S.R. Budihas, S.F.J. Le Grice SFJ, J.A. Beutler, A fluorescence-based high throughput screening assay for inhibitors of human immunodeficiency virus-1 reverse transcriptase-associated ribonuclease H activity, Anal. Biochem. 322 (2003) 33-39.

[19] M. Álvarez, T. Matomoros, L. Menéndez-Arias, Increased thermostability and fidelity of DNA synthesis of wild-type and mutant HIV-1 group O reversetranscriptases, J. Mol. Biol. 392 (2009) 872-884.

[20] S.J. Schultz, M. Zhang, J.J. Champoux, Sequence, distance, and accessibility are determinats of 5'-end-directed cleavages by retroviral RNases H, J. Biol. Chem. 281 (2006) 1943-1955.

[21] S.J. Schultz, M. Zhang, J.J. Champoux, Preferred sequence within a defined cleavage window specify DNA 3' end-directed cleavages by retroviral RNases H, J. Biol. Chem. 284 (2009) 32225-32238.

[22] P.L. Boyer, H.-Q. Gao, P. Frank, P.K. Clark, S.H. Hughes, The basic loop of the RNase H domain of MLV RT is important both for RNase H and for polymerase activity, Virology 282 (2001) 206-213.

[23] D. Lim, M. Orlova, S.P. Goff, Mutations of the RNase H C helix of the Moloney murine leukemia virus reverse transcriptase reveal defects in polypurine tract recognition, J. Virol. 76 (2002) 8360-8373.

[24] Z. Hostomsky, Z. Hostomska, G.O. Hudson, E.W. Moomaw, B.R. Nodes, Reconstitution in vitro of RNase H activity by using purified N-terminal and C-terminal domains of human immunodeficiency virus type 1 reverse transcriptase, Proc. Natl. Acad. Sci. USA 88 (1994) 1148-1152.

[25] J.L. Keck, S. Marqusee, Substitution of a highly basic helix/loop sequence in the RNase H domain of human immunodeficiency virus reverse transcriptase restored its Mn^{2+}-dependent RNase H activity, Proc. Natl. Acad. Sci. USA 92 (1995) 2740-2744.

[26] T. Tadokoro, S. Kanaya, Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechamism of the prokaryotic enzymes, FEBS J. 276 (2009) 1482-1493.

[27] S.J. Stahl, J.D. Kaufman, S. Vikić-Topić, R.J. Crouch, P.T. Wingfield, Construction of an enzymatically active ribonuclese H domain fo human immunodeficiency virus type 1 reverse transcriptase, Protein Eng. 7 (1994) 1103-1108.

[28] E.-D.Permanasari, K. Yasukawa, S. Kanaya, Enzymatic activities of RNase H domains of HIV-1 reverse transcriptase with substrate binding domains of bacterial RNase H1 and H2, Mol. Biotechnol. in press (2015)

[29] S. Kanaya, A. Kohara, Y. Miura, S. Sekiguchi, S. Iwai, H. Inoue, E. Ohtsuka, M. Ikehara, Identification of the amino acid residues involved in an active site of *Escherichia coli* ribonuclease H by site-directed mutagenesis, J. Biol. Chem. 265 (1990) 4615-4621.

[30] S. Kanaya, C. Katsuda-Nakai, M. Ikehara, Importance of the positive charge cluster in *Escherichia coli* ribonuclease H1 for the effective binding of the substrate, J. Biol. Chem. 266 (1991) 11621-11627.

[31] M. Nowotny, S.A. Gaidamakov, R. Ghirlando, S.M. Cerritelli, R.J. Crouch, W. Yang, Structure of human RNase H1 compexed with an RNA/DNA hybrid: insight into HIV reverse transcription, Mol. Cell. 28 (2007) 264-276.

Figure legends

Fig. 1. Expression plasmids for the RNase H domain. The structures of pET-RNaseH-WT, pET-RNaseH-ΔC1, and pET-RNaseH-ΔC2 are shown. The “498”, “671”, and asterisk indicate Ile498, Leu671, and the termination codon. The amino acid sequences of strep tag, HRV 3C protease recognition site, thrombin recognition site, and $(His)_{10}$ are underlined.

Fig. 2. SDS-PAGE analysis. CBB-stained 12.5% SDS-polyacryl amide gel is shown. (A) Active fractions of each purification stage for the wild-type RNase H (WT). Lanes: marker proteins (lane 1), soluble fractions of the total extracts (lane 2), active fractions of ion-exchange chromatography (lane 3), Ni^{2+} affinity chromatography (lane 4), and Strep-Tactin affinity chromatogaraphy (lane 5). (B) Purified enzyme preparations with N-terminal strep tag and C-terminal $(His)_{10}$ tag. Lanes: marker proteins (lane 1), WT (lane 2), RNaseH-ΔC1 (lane 3), and RNaseH-ΔC2 (lane 4). (C) Purified WT preparations. Lanes: marker proteins (lane 1), before removal of fused tags (lane 2), and after removale of fused tags (lane 3).

Fig. 3. UV and CD spectroscopic analysis of purified WT. (A) UV spectra. (B) CD spectra.

Fig. 4. RNase H activity of the RNase H domain using a fluorescent substrate. (A) Sequences of the RNA-DNA hybrid, R18/D18. (B) Fluorescence spectra of R18/D18 with excitation at 490 nm. The reaction was carried out with either of 50 nM HIV-1 RT, 1.8 μM WT, 1.8 μM ΔC1, and 1.8 μM ΔC2 for 30 min at pH 8.0 at 37ºC.

Fig. 5. RNase H activity of the RNase H domain using a radioactive substrate. (A) Sequences of the RNA-DNA hybrid, R25/D25. (B, C) Patterns of denatured PAGE. The enzyme concentrations were 30 nM for full-length MMLV RT, 1 μM for WT, and 10 μM for ΔC1 and ΔC2. The reaction times were 15 s and 10 min (B) and 15 s, 30 s, 1 min, 5 min, 15 min, 30 min, 1 h, and 2 h (C). The reaction was carried out with 1.8 μM WT, ΔC1, or ΔC2 for 30 min at pH 8.0 at 37ºC.

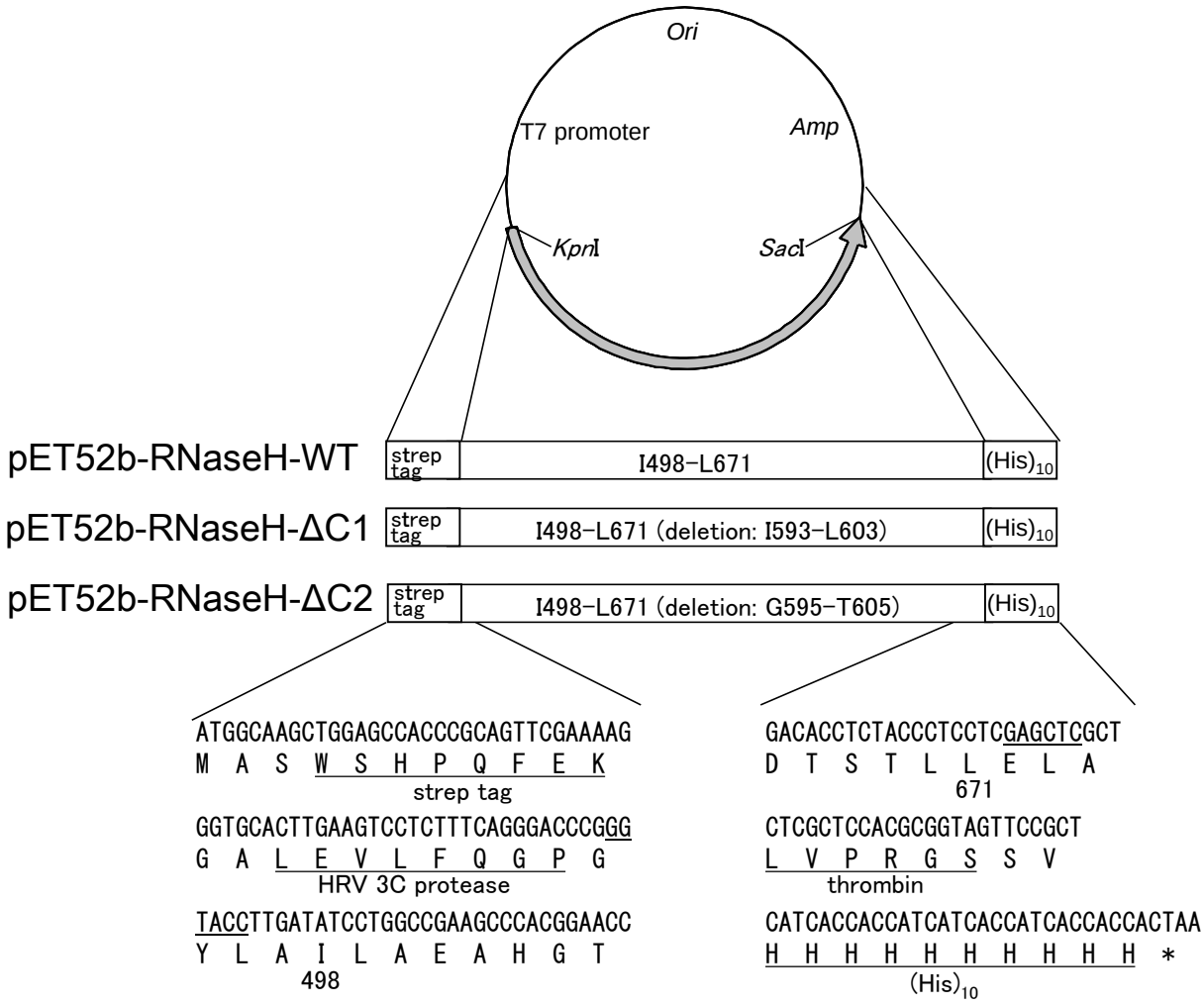
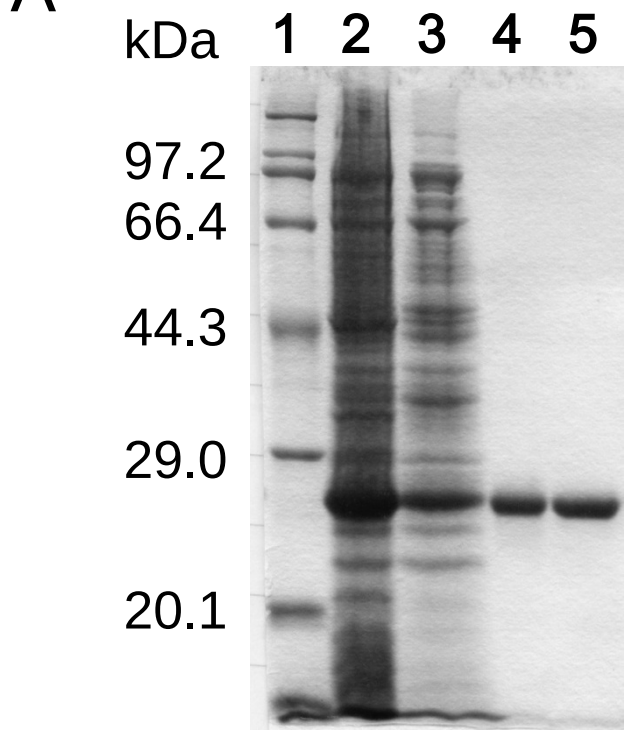
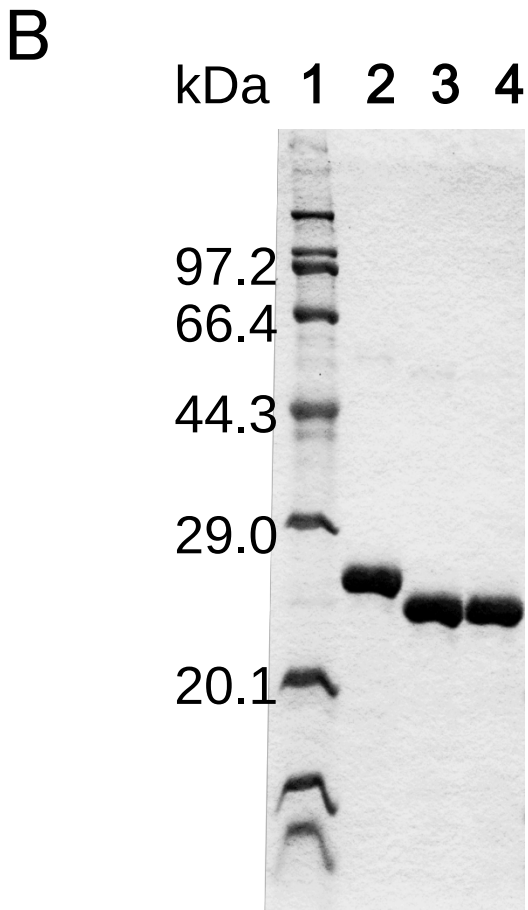


Fig. 1





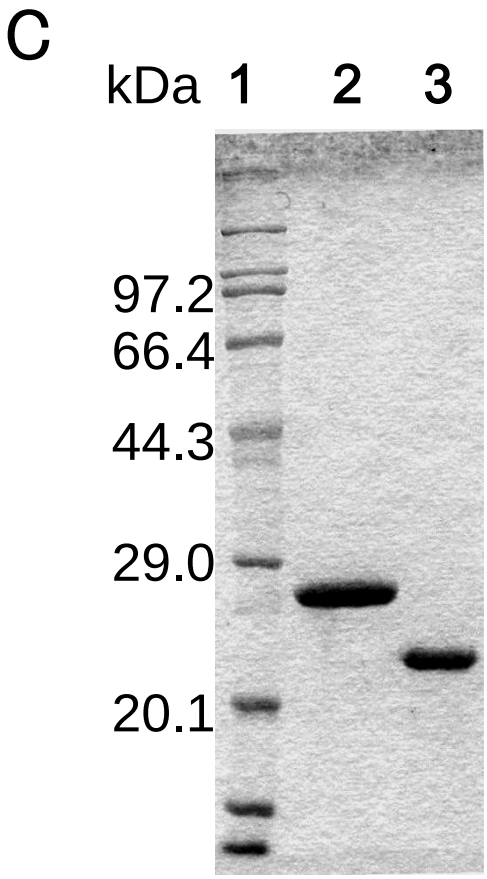
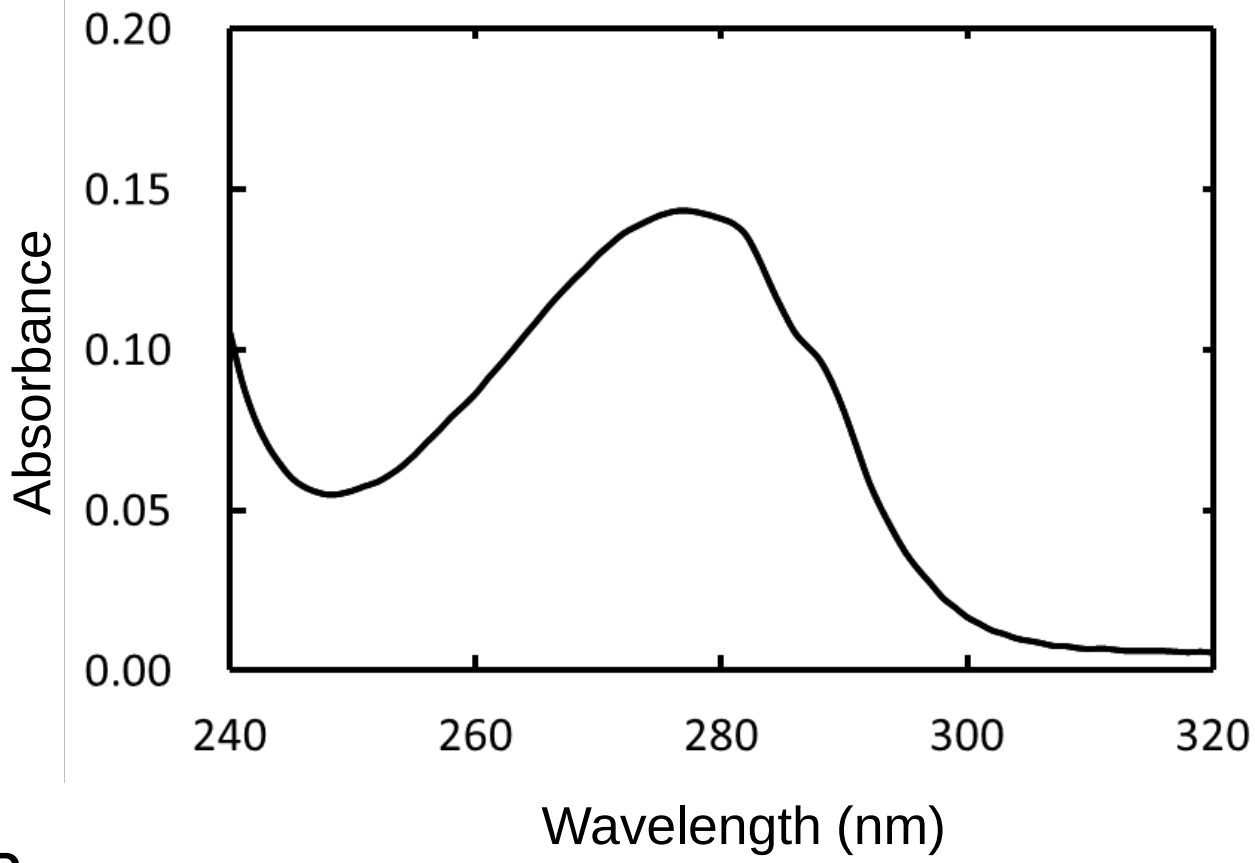


Fig. 2

A



B

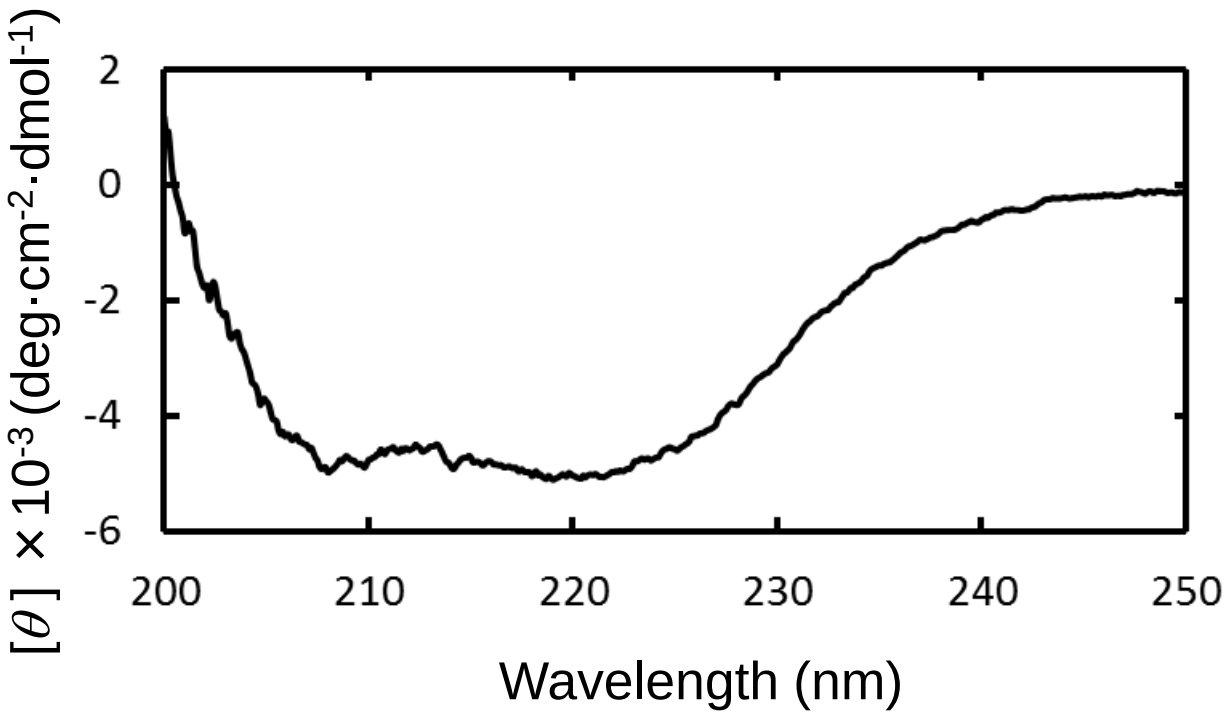


Fig. 3

A

R18: 5'-GAUCUGAGCCUGGGAGCU-fluorescein-3'
D18: 3'-CTAGACTCGGACCCTCGA-DABCYL-5'

B

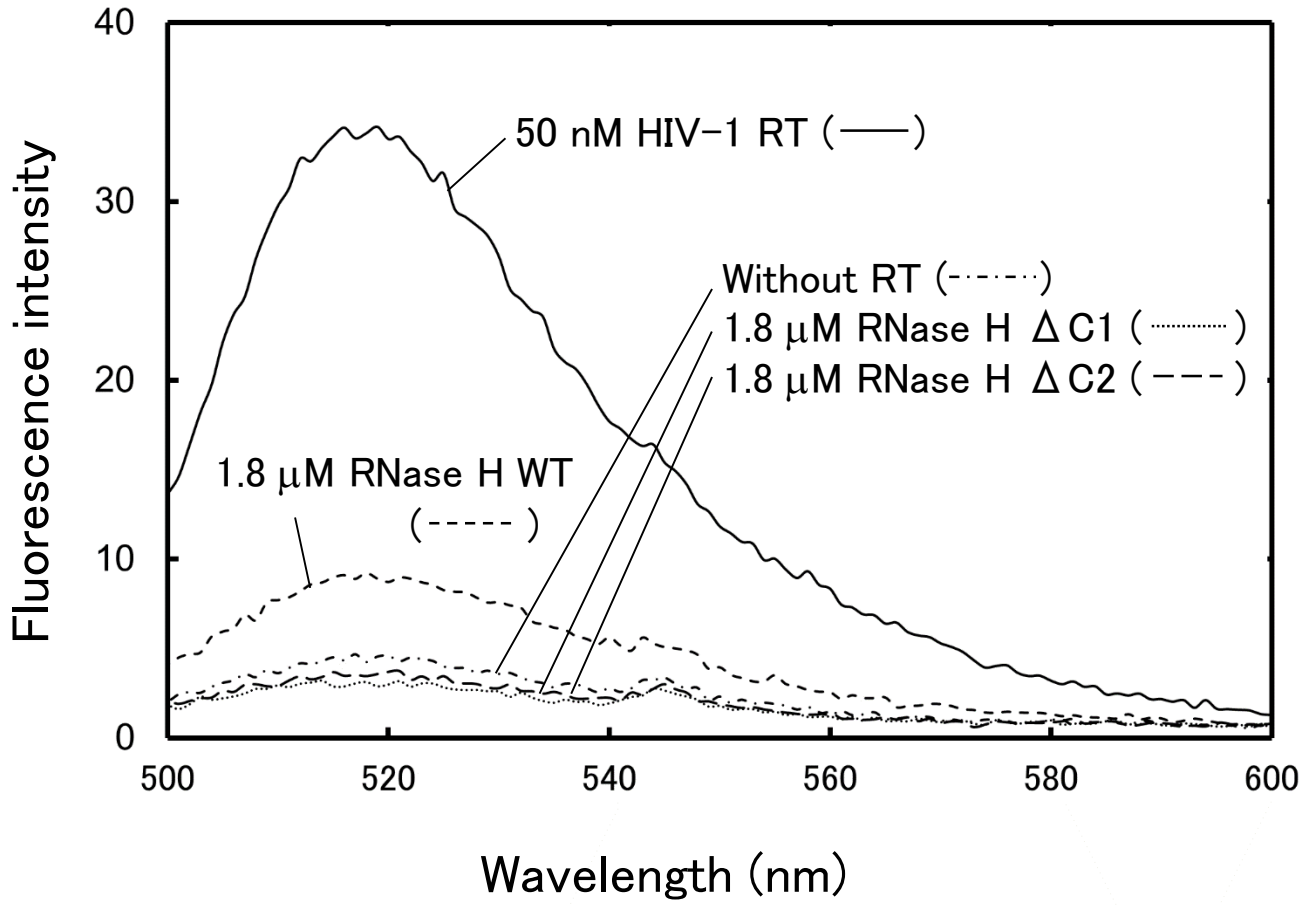


Fig. 4

R25: 5´-[^{32}P]AUGUAUAGCCCUACCAGCAUUCTGG-3´
D25: 3´- TACATATCGGGATGGTCGTAAGACC-5´

B

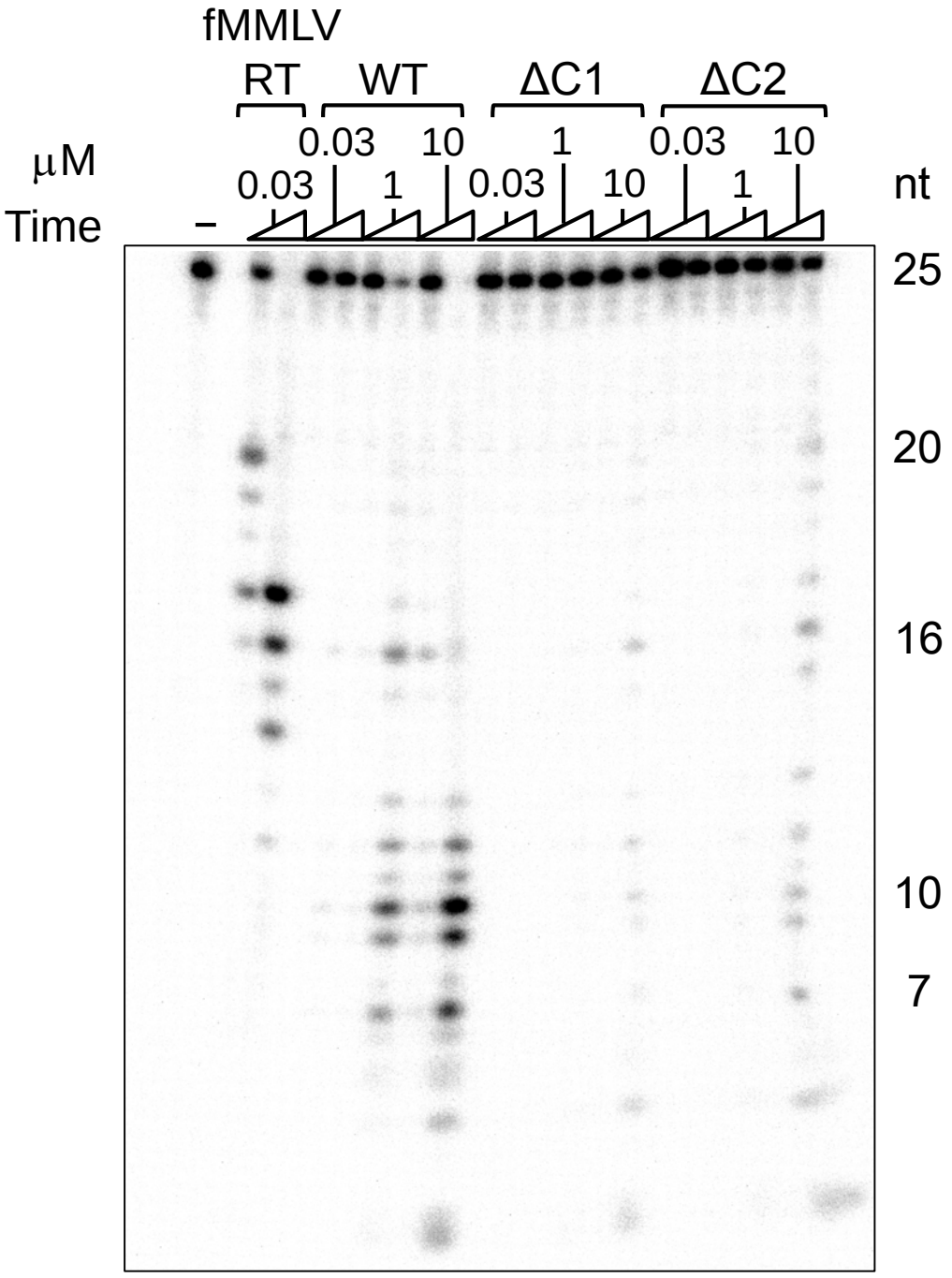


Fig. 5

C

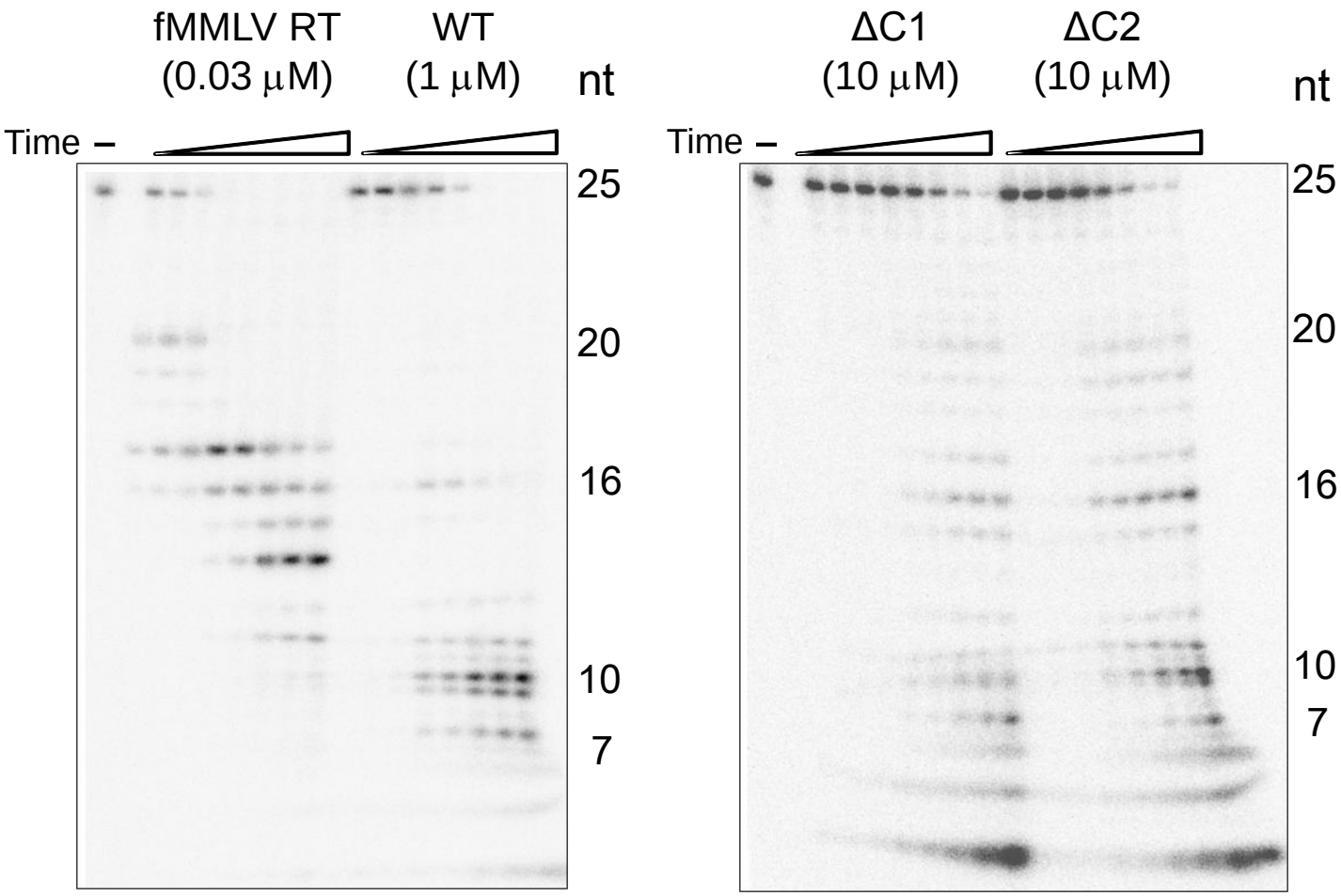


Fig. 5

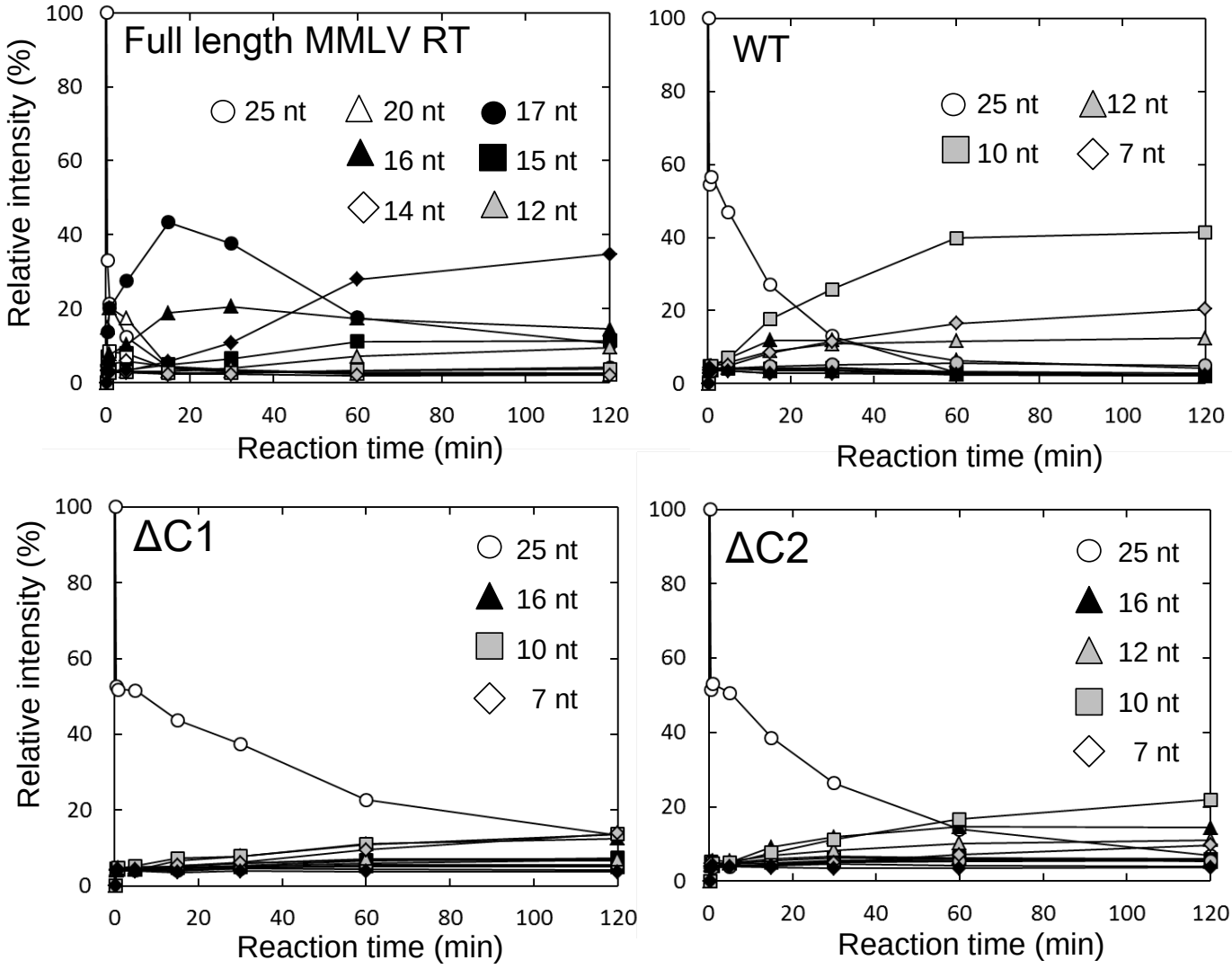


Fig. S1. Time course of the amounts of RNA fragments in the RNase H reaction. The result of the densitometry traces of the bands in Fig. 5C is shown. The relative intensity is defined as the ratio of the amount of each band to those of all 7−25-nt bands. Symbols for RNA fragments (nt): 25 (open circle), 20 (open triangle), 19 (open square), 18 (open diamond), 17 (solid circle), 16 (solid triangle), 15 (solid square), 14 (solid diamond), 13 (gray circle), 12 (gray triangle), 10 (gray square), and 7 (gray diamond).